Display awareness in subjective and objective video quality evaluation

Sylvain Tourancheau, Patrick Le Callet, Kjell Brunnström, Dominique Barba

DISPLAY AWARENESS IN SUBJECTIVE AND OBJECTIVE VIDEO QUALITY EVALUATION

Sylvain Tourancheau[1], Patrick Le Callet[1], Kjell Brunnström[2] and Dominique Barba[1]

(1) Université de Nantes, IRCCyN laboratory
rue Christian Pauc, 44306 Nantes, France
phone: +33 (0)240 683 047
patrick.lecallet@univ-nantes.fr

(2) Acreo AB, Video and Display Quality
Electrum 236, SE-164 40 Kista, Sweden
phone: +46 8 632 77 00
kjell.brunnstrom@acreo.se

ABSTRACT

Most of studies on video quality assessment are focused on impact of coding distortion or transmission error. In this paper, display is considered. Regarding technology, subjective experiments suggest that there are differences in term of quality between LCD and CRT. CRT provides better quality when viewing HD video content while it is LCD when viewing still colour images. One explanation of this behaviour is explained through motion blur. From a motion blur perception model, an efficient metric of the quality loss due to this effect on LCD is proposed. Finally, subjective experiments results using SD video content are consistent with a motion blur based model and point out that the effect of display technology is linked with video resolution.

1. INTRODUCTION

The incoming of the high-definition new visual experience at home have boosted new display technologies, since they enable the increase of screen size necessary to sense immersion, impact and immediacy as in a movie theatre [1]. For these reasons, these new displays, and particularly liquid crystal displays (LCD), will soon replace the old mature CRT technology.

Liquid crystal displays have many differences with CRT displays. Subjective preference tests between this two types of displays have highlighted a high preference for CRT displays concerning moving pictures [2]. Many defects have been counted by viewers, such as colour differences, degradations in dark areas and de-interlacing artifacts for interlaced sequences. But among all these defects, motion blur seems to be the most annoying one, particularly in sequences with significant movements. On the other hand, CRT displays suffer from several shortcomings too. Flickering can be annoying in certain conditions, and the small luminance range can led to flat pictures with dirty colours.

In the recent years, subjective and objective quality assessment becomes a research topic of interest. Activities of the Video Quality Experts Group (VQEG) are a good example of this interest. Previous works [3] and work in progress [4] are mainly related to coding or transmitting purpose at a given resolution, e.g. coding artifacts and transmission errors. Considering the whole chain, quality assessment should be able to manage dependency to other technology issues.

In this paper, display is studied. As a consequence, only the high part of the quality range is considered, using video and pictures with no (uncompressed) or very few coding distortions. In order to know the impact of distortions due to flat panel displays technology with respect to the old mature CRT technology, subjective quality assessment of HDTV sequences is performed both on LCD and CRT displays.

Assuming that LCD motion blur is the most annoying artifacts when displaying moving pictures on LCD, its perception is described and its magnitude measured. This leads to the design of an objective metric which enables the prediction of the perceived quality on LCD with respect to the CRT one. Such a metric could be used in order to evaluate LCD improvements introduced by manufacturers to reduce technology artifacts.

In the last section, the impact of LCD technology on perceived quality regarding the video resolution is explored with subjective experiments on SD sequences.

2. SUBJECTIVE QUALITY ASSESSMENT TESTS

2.1 Tests conditions and equipment

Subjective quality assessment tests have been performed in a specific showroom, with lighting conditions and display parameters precisely measured and adjusted according to BT.500-11 and BT.710-4 ITU recommendations. Tests have been conducted both on a HDTV CRT display JVC DT-V 1910CG and then on a HDTV LCD Philips T370 HW01. Surrounding conditions and display parameters for each screen are presented in Tables 1 and 2. Viewing distances have been set to $3H$, where H is the height of the displayed pictures.

Background luminance	7 cd/m^2
Chromaticity of background	D_{65}
Picture height (H)	20.5 cm
Viewing distance	61.5 cm ($3H$)
Display black luminance	0.531 cd/m^2
Display peak luminance	70.9 cd/m^2

Table 1: Viewing conditions and display parameters for the CRT display JVC DT-V 1910CG.

2.2 Protocol

These tests have been performed with sequences and pictures of rather good to excellent quality. As a consequence, the used protocol should enable the quality discrimination. A well known stable method for this purpose is the SAMVIQ protocol [5], developed by France Telecom RD and standardised by the EBU and the ITU.

SAMVIQ is multi stimuli continuous quality scale protocol. With this procedure, the observers can compare the

Background luminance	35 cd/m^2
Chromaticity of background	D_{65}
Pictures height (H)	46 cm
Viewing distance	138 cm ($3H$)
Display black luminance	0.641 cd/m^2
Display peak luminance	471 cd/m^2

Table 2: Viewing conditions and display parameters for the LC display Philips T370 HW01.

quality of processed sequences (resp. pictures) both between them and with the quality of the explicit reference sequence (resp. picture). This leads to precise and reliable measures of the quality [6]. Notation scale is continuous, each score can take a value between 0 and 100.

2.3 Observers

Observers were mostly male students between 20 and 25. All were familiar with television and cinema but not with HDTV. Acuity and colour blindness have been checked for each observer, respectively with Monoyer's plates and Ishihara colour vision test. Observers with at least on error in Ishihara test or with an acuity less than 9/10 are rejected.

After the tests have been completed by all the observers, a rejection technique from the EBU [5] is applied. This process verifies the level of consistency of the scores of one observer according to mean score of all observers. Following the application of this rejection process, 15 valid subjects should be retained at minimum.

3. VIDEO QUALITY ASSESSMENT

3.1 Material

In order to measure the difference of quality between the two types of displays for moving pictures, nine 1080i sequences with significant movements have been chosen. These videos have been supplied by European broadcasters SVT and Euro1080.

Each of them contains 250 frames which corresponds to a 10-second duration. Each reference (uncompressed) has been distorted with the H.264 reference coder at seven different bit-rates in a range of quality from . Sequences are received in 1080i format by the two displays. They're displayed in interlaced format on the CRT but not on the LCD which de-interlaces them since the flat panel matrix are only able to display progressive format.

3.2 Results

Mean opinion scores (MOS) of the observers for the ten reference sequences and for the two types of displays are presented in Table 3. ΔMOS is the difference of MOS from CRT and LCD:

$$\Delta\text{MOS} = \text{MOS CRT} - \text{MOS LCD} \tag{1}$$

The perceived quality of moving pictures displayed on LCD is globally lower than the perceived quality of moving pictures displayed on CRT display. It's interesting to notice that this loss of quality is quite important for sequences with quick movements such as *Concert*, *Parkrun*, *Foot* and *Voile*.

Séquence	MOS CRT	MOS LCD	ΔMOS
VOILE	83.9	77.7	6.2
FOOT	82.8	76.3	6.5
CONCERT	84.5	73.8	10.7
SHOW	82.9	75.3	7.6
CREDITS	83.1	79.1	4.0
MOBCAL	81.4	81.0	0.4
PARKRUN	87.6	80.2	7.4
SHIELDS	86.7	78.2	8.5
STOCKHOLM	86.1	82.3	3.8

Table 3: Mean opinion scores by sequences and displays.

This loss of quality on LCD seems to be related to the quantity and/or the fastness of the movements in the sequence. To validate this hypothesis, it has been decided to conduct the same experiment with still pictures.

4. STILL PICTURES QUALITY ASSESSMENT

4.1 Material

Five images have been chosen in order to measure the difference of quality between the two types of screens for still images. They contain specific contents such as natural textures, flesh colours, oriented contours, water reflection, characters, etc. Each of them has been distorted with two types of process: jpeg compression and down-scaling/up-scaling filtering. Pictures are displayed in interlaced format on the CRT and in progressive format on the LCD in order to repeat exactly the same conditions as those of the video quality assessment.

For this test, the same group of observers has been used for the two displays. The group has been split in two parts: the observers of the first part have passed the test on CRT first, the observers of the second part have passed the test on LCD first.

4.2 Results

MOS of the observers for the five reference pictures and for the two displays are presented in Table 4.

Séquence	MOS CRT	MOS LCD	ΔMOS
FOOTBALL	66.3	79.4	-13.1
HAND	73.6	80.3	-6.7
HOUSE	51.8	81.8	-30.0
LANDSCAPE	73.5	78.7	-5.2
MAP	51.4	84.4	-33.0

Table 4: Mean opinion scores by pictures and displays.

It can be observed that for still pictures the quality on LCD is globally preferred. For the pictures *House* and *Map* the difference between the two types of displays is largely in favour of LCD (with a ΔMOS of about a third of the quality scale). This can be explained by the presence of horizontally fine oriented contours which make the flickering of the CRT display more noticeable.

Overall, shortcomings of CRT displays such as flickering and limited range of luminance seems to lead to a lower feeling of natural and sense of immersion. LCD is brighter, vivid and colourful and the perceived quality of still pictures is clearly higher on it (ΔMOS_{mean} = -17.6). However, with the exact same viewing conditions and displays parameters, the perceived quality of moving pictures is higher on CRT (ΔMOS_{mean} = 6.4). It's assumed that this difference must be due to moving artifacts such as LCD motion blur which are not present with still pictures.

5. LCD MOTION BLUR

Results described in the previous part lead to the statement that the excellent perceived quality on LCD with still pictures is strongly reduced with moving pictures. Moving artifacts due to LCD technology, and particularly LCD motion blur, seem to be responsible for this loss of quality in video.

In this part, LCD motion blur is described. The perception of this motion blur is then measured and a model of perception is proposed. Finally, this perception model is used to design an objective metric which enables the prediction of the loss of quality on LCD with respect the perceived quality on CRT.

5.1 Description

LCD motion blur has been widely studied in recent works [7, 8, 9]. It's mainly caused by the hold-type LCD's displaying method: the light intensity is maintained on the screen for the duration of the frame, whereas on CRT light intensity is a pulse which fades over the frame duration.

The main difference happens when the eyes of the observer are tracking a moving object on the LCD screen: for a given frame, the picture is sustained on the screen while the eyes are still moving slightly anticipating the movement of the object. Edges of this object are displaced on the retina resulting in a blur [10].

5.2 Model of motion blur perception

In order to measure the relation between the motion velocity and the magnitude of perceived blur in the simple case of scrolling bars, psycho-physics measurements have been designed [11]. Results of these experiments are presented in Figure 1, they lead to the following linear relation :

$$W = aV; \tag{2}$$

The width W (in pixels) of motion blur that appears on the edges of a moving object is proportional to its velocity V (in pixels per frame).

Pan *et al.* have developed a theoretical model of LCD motion blur perception [8] and they obtain the same relation. Their model permits to identify the parameter a, which depends on the temporal function of the display.

5.3 Prediction of ΔMOS based on motion blur

Since subjective assessment on HD video content provides a value of ΔMOS between CRT and LCD per sequence, an objective metric is developed in order to predict this quality loss ΔMOS_p from the motion blur perception model. This metric is made in several steps. First, a spatio-temporal classification is done in two passes. First pass is corresponding to a block based motion estimation that leads to the construction of tubes which are the sets of blocks positions along the direction of motion. Second pass is the classification of each tube according to its spatial content. Since motion blur is

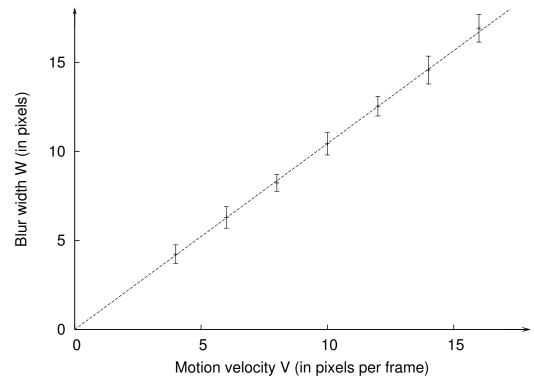

Figure 1: Perceived blur width W as a function of motion velocity V.

only visible at sufficient contrast [12], only tubes categorised as textures and edges are selected. An average motion vector is computed from all the vectors of the remaining tubes. Norm of this global vector is used to compute the width of perceived motion blur according to Equation 2. This value W is an indicator of the average magnitude of perceived blur along the sequence. Finally, ΔMOS_p is computed from a function of W. This function is non linear since there is no influence on perceived quality below a threshold of W, and the quality difference saturates for high values (cf. Figure 2).

An estimation of the subjective quality scores on LCD from the subjective quality scores on CRT can be made using the following relation:

$$\text{MOS LCD}_{est} = \text{MOS CRT} - \Delta\text{MOS}_p. \tag{3}$$

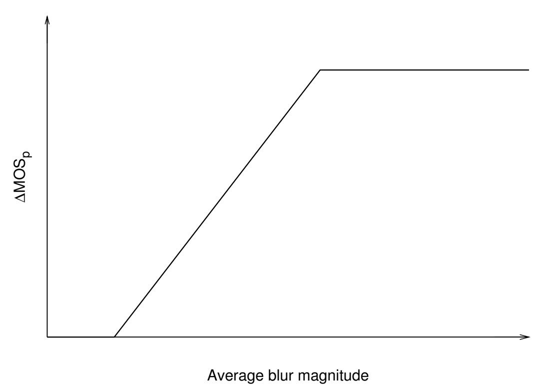

Figure 2: Prediction of the loss of quality ΔMOS_p from the average blur magnitude.

The quality of the model can be measured by the linear correlation coefficient (CC) and the root mean square error (RMSE) between estimated LCD scores and actual LCD scores. Values of 0.958 for CC and 1.30 for RMSE are obtained. These performances can be compared with those obtain comparing results between CRT and LCD: 0.241 for CC and 6.75 for RMSE.

Therefore, the proposed prediction model is able to evaluate solutions to LCD motion blur reduction from obtained quality on CRT.

6. IMPACT OF LCD MOTION BLUR WITH RESPECT TO THE DISPLAY RESOLUTION

As it has been shown in the previous section, a high correlation exists between the magnitude of perceived blur in a sequence and the loss of quality observed on LCD when displaying this sequence. Furthermore, the magnitude of perceived blur depends on display characteristics and resolution. More precisely in Equation 2, the velocity V of motion would be less important if the resolution is lower. As a consequence, the perceived blur would be smaller and should be less annoying as the resolution decreases.

6.1 Standard definition video quality assessment

In order to assess the impact of LCD motion blur relatively to the resolution, the same video quality assessment tests have been realised with sequences at a standard definition (SD). Four sequences of the previous experiment have been chosen and reduced to SD resolution by computing the HD versions through a half-band filtering followed by a down-sampling by a factor of 2 (both along horizontal and vertical directions). This processing is performed on each field of the interlaced 1080i sequences, with the advantage that it doesn't necessitate any interpolation.

Tests have been led both on the CRT display and then on the LCD. Viewing conditions and display parameters are the same as those described in Section 2.1, except for the viewing distances which have been set to six times the pictures'height ($6H$). SD sequences have been displayed inserted in a HD resolution grey level sequence in order to suit the displays native HD resolution.

Each reference (uncompressed) has been distorted with the H.264 reference coder at seven different bit-rates. The impact of technology on the perceived quality is assessed. As a result, only the sequences with few or no coding distortions are taken in account: the mean of MOS for sequences whith good to excellent quality is computed.

6.2 Results

Table 5 shows the mean of MOS for good-to-excellent quality HD sequences, on the two displays, and ΔMOS is computed. Same results for SD sequences is shown in Table 6. As expected, the loss of quality on LCD is strongly smaller in SD resolution with respect to HD resolution. Moreover, the loss of quality on SD sequences is not significant with regards to intervals of confidence.

As explained before, the perception of LCD motion blur closely depends on the display resolution. The quantity of perceived blur linearly depends on the velocity of motion which is twice smaller in SD than in HD (since the resolution is divided by two). As a result, the perceived motion blur should be less annoying in SD and the advantages of LCD such as colourfulness and large luminance range seem to tower over this artifact leading to a better global quality on LCD display.

Séquence	MOS CRT	MOS LCD	ΔMOS
MOBCAL	79.5	71.9	7.6
PARKRUN	83.0	70.7	12.3
SHIELDS	81.4	67.9	13.5
STOCKHOLM	81.6	73.0	8.7

Table 5: Mean of the MOS for good-to-excellent quality sequences in HD resolution, on CRT and LCD.

Séquence	MOS CRT	MOS LCD	ΔMOS
MOBCAL	71.6	68.0	3.6
PARKRUN	77.6	72.9	4.7
SHIELDS	75.5	73.5	2.0
STOCKHOLM	75.7	73.2	2.5

Table 6: Mean of the MOS for good-to-excellent quality sequences in SD resolution, on CRT and LCD.

7. CONCLUSION

Subjective quality evaluation of HD moving pictures have shown that the perceived quality is better on CRT display than on LCD. This loss of quality ΔMOS seems to be due to the flat panel technology. Actually, new artifacts such as motion blur are very annoying with quick movements. Benefits of LCD, for instance colourfulness and larger luminance range, have been highlighted with subjective assessment of still pictures: they lead to a better perceived quality on LCD than on CRT. However, in video they don't achieve to compensate the loss of quality due to moving artifacts.

LCD motion blur have been studied and a mathematical model is used to measure its magnitude as a function of the quantity of movements. This magnitude depends on the display resolution in the same way as the velocity of moving objects. A high correlation has been highlighted between this motion blur magnitude and the loss of quality on LCD which enables the prediction of the loss of quality ΔMOS between CRT and LCD.

Consequently, the loss of quality ΔMOS due to LCD motion blur would depend on the display resolution and should be weaker with lower resolutions than HD. Same video quality assessment tests with SD sequences have confirmed this. At a lower resolution, the moving artifacts due to LCD technology are less annoying. Benefits of LCD such as colourfulness and larger luminance range seem to tower over this defects since the perceived quality is better on LCD in SD resolution.

In this paper, it's shown that the new LCD technology leads to new shortcomings when displaying moving pictures. These new artifacts are not significant at low resolutions, which validate the use of LCD for subjective video quality assessment in the Multimedia Testplan of the VQEG [4]. However, when increasing the resolution, artifacts due to flat panel technology become more annoying and have an important impact on the perceived quality. It comes that the subjective video quality assessment at high resolution (HDTV for example) should be led very carefully on LCD since a significant part of perceived distortions could be due to the display.

REFERENCES

[1] Takashi Fujio, "Future broadcasting and high definition television," NHK technical monograph, NHK, June 1982.

[2] ITU, "Report on results of comparative subjective picture quality assessment test between CRT and LCD," Questions ITU-R 95/6, 102/6, International Telecommunication Union - Radiocommunication Study Groups, August 2005.

[3] VQEG, "Final report from the video quality experts group on the validation of objective models of video quality assessment," Tech. Rep., VQEG, 2003, http://www.its.bldrdoc.gov/vqeg/projects/frtv_phaseII/downloads/VQEGII_Final_Report.pdf.

[4] VQEG, "Multimedia group test plan," Tech. Rep. Draft version 1.16, VQEG, February 2007, http://www.its.bldrdoc.gov/vqeg/projects/multimedia/MM_new_testplan_v1.16_changes_accepted.doc.

[5] EBU, "SAMVIQ – Subjective assessment methodology for video quality," Tech. Rep., European Broadcasting Union, 2003.

[6] Jean-Louis Blin, "New quality evaluation method suited to multimedia context: Samviq," in *Proceedings of the Second International Workshop on Video Processing and Quality Metrics, VPQM'06*, Scottsdale, January 2006.

[7] Michiel A. Klompenhouwer, "The temporal MTF of displays and related video signal processing," in *IEEE International Conference on Image Processing, 2005. ICIP 2005.*, September 2005, vol. 2, pp. 13–16.

[8] Hao Pan, Xiao-Fan Feng, and Scott Daly, "LCD motion blur modeling and analysis," in *IEEE International Conference on Image Processing, 2005. ICIP 2005.*, September 2005, vol. 2, pp. 21–24.

[9] Xiao-Fan Feng, "LCD motion-blur analysis, perception, and reduction using synchronized backlight flashing," in *Proceedings of the SPIE Conf. Human Vision and Electronic Imaging XI*. Electronic Imaging 2006, Janvier 2006, vol. 6057.

[10] Taiichiro Kurita, "Moving picture quality improvement for hold-type AM-LCDs," *SID Symposium Digest of Technical Papers*, vol. 32, no. 1, pp. 986–989, June 2001.

[11] Sylvain Tourancheau, Patrick Le Callet, and Dominique Barba, "Prediction of perceived quality differences between CRT and LCD displays based on motion blur," in *Proceedings of the Third International Workshop on Video Processing and Quality Metrics for Consumer Electronics, VPQM2007*, Scottsdale, 2007.

[12] Justin Laird, Mitchell Rosen, Jeff Pelz, Ethan Montag, and Scott Daly, "Spatio-velocity CSF as a function of retinal velocity using unstabilized stimuli," in *Proceedings of the SPIE Conf. Human Vision and Electronic Imaging XI*. Electronic Imaging 2006, Janvier 2006, vol. 6057.